Polytopic Graph of Latent Relations

A Multiscale Structure Model for Music Segments

Corentin Louboutin and Frédéric Bimbot

IRISA (CNRS - UMR 6074 & Université Rennes 1)
Campus de Beaulieu, 263 avenue du Général Leclerc, 35042 RENNES cedex, France
corentin.louboutin@irisa.fr frederic.bimbot@irisa.fr

Abstract. Musical relations and dependencies between events within a musical passage may be better explained as a graph rather than in a sequential framework. This article develops a multiscale structure model for music segments, called Polytopic Graph of Latent Relations (PGLR) as a way to describe nested systems of latent dependencies within the musical flow. The approach is presented conceptually and algorithmically, together with an extensive evaluation on a large set of chord sequences from a corpus of pop songs. Our results illustrate the efficiency of the proposed model in capturing structural information within such data.

1 Presentation

It is quite common sense that listeners do not perceive music only as a mere sequence of sounds, nor composers conceive their works as such. Music is essentially the result of patterns whose inner organization and mutual relationships participate to the overall structure of the musical content, at different time-scales simultaneously.

What is exactly music structure remains an open scientific question. This article is a contribution towards one particular aspect of music structure: it proposes and investigates a multiscale model of the inner organization of musical segments, which we call Polytopic Graph of Latent Relations (PGLR).

The musical content observed at a given instant t within a music segment obviously tends to share privileged relationships with its immediate past, hence the sequential perception of the music flow. But music content at instant t also relates with distant events which have occurred in the longer term past, especially at instants which are metrically homologous to t, in previous bars, motifs, phrases, etc. This is particularly evident in strongly "patterned" music, such as pop music, where recurrence and regularity play a central role in the design of cyclic musical repetitions, anticipations and surprises. But it is also discernable in a number of other music genres, which rely abundantly on all sorts of multiscale similarities, progressions, expectations and denials.

To overcome the limitations of purely sequential models in music content descriptions, hierarchical models are often resorted to, in order to provide a representation framework for the grouping structure of a musical passage. The most famous hierarchical approach is undoubtedly the Generative Theory of Tonal Music (GTTM) by Lerdahl and Jackendoff (8), which has been for many years a source of inspiration for a wide variety of work in music structure modeling. However, hierarchical approaches such as GTTM rely axiomatically on an adjacency hypothesis, under which the grouping of elements into a higher level object is strictly limited to neighbouring units.

In this work, we develop a different view as regards the structural association of elements forming music segments. We describe the "web" of musical elements as a *Polytopic Graph of Latent Relations* (PGLR) which models relationships developing predominantly between homologous elements within the metrical grid.

For most segments of 2^n events, the PGLR lives on an n-dimensional cube (square, cube, tesseract, etc...), n being the number of scales considered simultaneously in the multiscale model. By extension, the PGLR can be generalized to a more or less regular n-polytope.

Each vertex in the polytope corresponds to a low-scale musical element, each edge represents a relationship between two vertices and each face forms an elementary system of relationships. In addition, one variant of the proposed model views the last vertex in each elementary system as the denied realization of a (virtual) expected element, itself resulting from the implication triggered by the combination of former relationships within the system.

The estimation of the PGLR structure of a musical segment can be obtained computationally as the joint estimation of:

1. the description of the polytope (as a more or less regular n-polytope),
2. the nesting configuration of the graph over the polytope, reflecting the flow of dependencies and interactions between the elementary implication systems within the musical segment (this flow being assumed to be causal),
3. the set of relations between the nodes of the graph, with potentially multiple possibilities which need to disambiguated (hence the "latent" nature of the relations, as they are not actually observed).

In this paper, the shape of the polytope is assumed to be a tesseract (4-cube) and we focus our study on the modeling of meter-synchronous chord sequences of 16 chords. However, the general framework encompassed by PGLR is potentially applicable to many other musical dimensions (rhythm, melody, etc...) as soon as relevant latent relations can be defined.

In Section 2, we introduce the main concepts and formalism related to the model. Section 3 covers computational aspects of the approach, including optimality criteria and algorithmic design. In section 4, we present a series of experimental results which assess the advantages of the PGLR model. We conclude with perspectives outlined by the proposed approached.

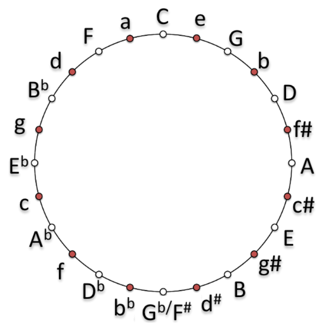

Fig. 1: Triads: circle of thirds.

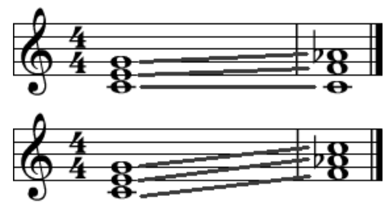

Fig. 2: Two possible transports between C and Fm.

2 Concepts and Formalism

2.1 Chord Representation and Relations

Strictly speaking, a chord, in music, is any harmonic set of notes (or "pitches") that are heard as if sounding simultaneously. However, in tonal western music, chords are more generally conceived as sets of *pitch classes* supporting the local harmonic groundplan of the music. In particular, chords play a strong role in the accompaniment of the melody in pop songs. The most frequently encountered chords are triads (i.e. sets of three pitch classes), with a predominance of major and minor triads. More sophisticated chords contain combinations of 4 pitch classes or even more.

Chords can be represented in various ways. In this article, we consider two types of representations: (*i*) the complete set of pitch classes forming the chord (PC description) and (*ii*) the tabular notation of the major or minor triadic reduction of the chord (TR description). Assuming 4 or 5 pitch classes per chord, this leads to potentially several hundreds of different PC descriptions (much less in practice), but only 24 distinct TRs.

A number of formalisms exists to describe chord relations, either in classical musicology (through chromatic relations or via the circle of fifths) or in the framework of more recent theories, in particular Wietzmann regions (21) or neo-Riemannian theory (3). Tymockzo (18; 19) also proposes a model based on combinations of chromatic and scalar transpositions.

Depending on the formalism under consideration, the property of uniqueness of the relation between two chords may or may not be satisfied.

Triad Circles We call *triad circle* any circular arrangement of triads aimed at reflecting some proximity relationship between triads along its circumference. The circle of thirds is formed by alternating major and minor triads with neighbouring triads sharing two common pitch classes. The circle is shown on Fig. 1.

This representation provides a way to express the relationship between two TRs – in a unique way – as the angular displacement around the circle. Alternatively, the chromatic circle is arranged according to a chromatic progression (not represented on Fig. 1).

Optimal Transport If two chords X and Y are represented as a set of pitch classes x_i and y_j, the set of *transports* between X and Y can be defined as:

$$T = \{t_k = (x_{i_k}, y_{j_k}) \mid x_{i_k} \in X,\ y_{j_k} \in Y\} \tag{1}$$

that is, pairs of notes across the two chords indexed by an integer k which represents a virtual mapping between their respective pitch classes. This is a simplified model that can be used to represent "voices" in chord sequences. We consider complete transports, i.e. each note is associated to at least one voice. Examples of transports are given on Fig. 2.

The *cost* of a transport is defined as the sum of the costs associated with each pair of notes in the transport: $|T| = \sum_{(x,y)\in T} |d(x,y)|$. In this paper we use two types of distances:

- the *chromatic distance* (or smoothness) (9; 16; 3), which is the shortest displacement in semitones from pitch class x to pitch class y. In Fig. 2 the first transport is minimal for the chromatic distance (cost equal to 2).
- the *harmonic distance*, where the displacement is considered on the circle of fifths instead of the chromatic scale. In Fig. 2, the second transport is minimal for the harmonic distance (cost equal to 6).

2.2 Systemic Organization

Based on the hypothesis that the relations between musical elements in a segment are not necessarily sequential, the System & Contrast (S&C) model has been recently formalized (1) as a generalization and an extension of Narmour's Implication-Realization model (14). Its applicability to various music genres for multidimensional and multiscale music analysis has been explored in (4) and algorithmically implemented in an early version as "Minimal Transport Graphs" (10).

The S&C model primarily assumes that relations between 4 elements in a musical segment $x_0\ x_1\ x_2\ x_3$ can be viewed as relying on a matrix-based *system* of relations in reference to the first element x_0 (the *primer*), which thus plays the role of a common *antecedent* to all other elements in the system. This is the basic principle that enables the joint modeling of two timescales simultaneously.

Moreover, in the S&C approach, it is further assumed that latent relationships $x_1 = f(x_0)$ and $x_2 = g(x_0)$ trigger a process of implication:

$$x_0\ f(x_0)\ g(x_0) \stackrel{implies}{\Longrightarrow} g(f(x_0)) = \hat{x}_3$$

Table 1: Antecedent function for the various models.

Sequential	Systemic	System & Contrast
$\phi_{Seq}(x_i) = x_{i-1}$	$\phi_{Sys}(x_i) = x_0$	$\phi_{S\&C}(x_i) = \begin{cases} x_0 & \text{if } i = 1, 2 \\ g(f(x_0)) & \text{if } i = 3 \end{cases}$

Virtual element $\hat{x}_3$ may be more or less strongly denied by a *contrast*: $x_3 \neq \hat{x}_3$, which creates a potential closure to the segment.

As depicted in Table 1, sequential, systemic and S&C models studied in this article are all first-order models which assume different antecedent functions, Φ, between the elements forming a musical segment. It is worth noting that the antecedent function summarizes the entire history of x_i into a single element.

2.3 Polytopic Representation and Nested Configurations

Polytopic Representation Elementary systems of 4 elements, as described in the previous section, can further be used to describe longer sequences of musical events. In particular, sequences of 2^n elements can be arranged as an n-dimensional cube, within which each face potentially forms a S&C at time instants that share specific relationships in the metrical grid.

For instance, a sequence of 16 chords can be divided into four sequences of four successive chords, each of them being described as separate systems. Then, these four S&Cs, taken as elementary objects, can be related by forming an upper-scale S&C, linking the four primers of the 4 lower-scale S&Cs. Fig. 3 represents such a description projected on a tesseract, in the case of the chord sequence from the chorus section of *Master Blaster* by Stevie Wonder:

Cm Cm Cm Bb   Ab Ab Ab Gm   F F F F   Cm Cm Bb Bb

System Nesting However, depending on the sequence, other arrangements of the systems may prevail. If we now consider the following example:

Bm Bm A A   G Em Bm Bm   Bm Bm A A   G Em Bm Bm

a different configuration appears to be more efficient to explain the sequence with a multiscale model. In fact, grouping chords $[0, 1, 8, 9]$, $[2, 3, 10, 11]$, $[4, 5, 12, 13]$ and $[6, 7, 14, 15]$, and then relating these four faces of the polytope by an upper-scale system $[0, 2, 4, 6]$ leads to a less complex (and therefore more economical) description of the relations between the data within the systems. This nesting configuration is called P^* in the rest of the paper and is distinct from the configuration considered in the first example, P_0, where the upper-scale system

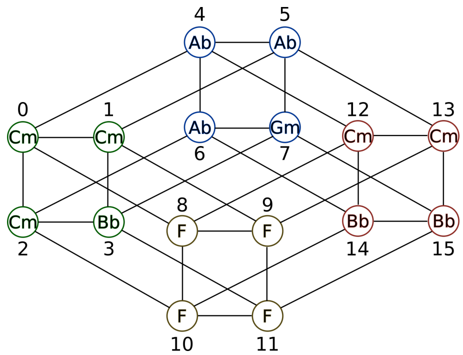

Fig. 3: Polytopic representation of the chord sequence taken from *Master Blaster* by Stevie Wonder.

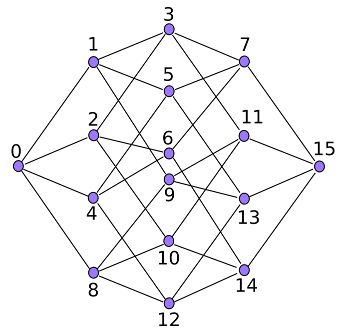

Fig. 4: Tesseract where elements of the same depth are aligned vertically.

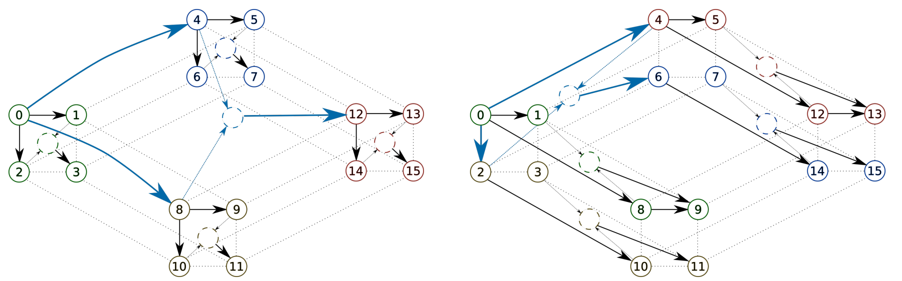

Fig. 5: Representations of the relations used by a multiscale analysis of a sequence of 16 events projected on a tesseract: P_0 (left), P^* (right).

$[0, 4, 8, 12]$ links four lower-scale nested systems $[4k + j]_{0 \leq j < 4}$ for $0 \leq k < 4$. Fig. 5 illustrates these two configurations.

Therefore, multiscale polytopic descriptions involve different possible flows of dependencies and interactions between systems, which correspond to distinct *nesting configurations*. A nesting configuration is characterized by its corresponding antecedent function (as defined in Sect. 2.2). We furthermore assume that nesting configurations must respect a causality principle: that is, the antecedent of any element in a system must have been observed before that element. This leads to a partial order between elements in the tesseract, as depicted on Fig. 4.

Static Configurations Among all possible ways to construct nested configurations, an interesting subset consists in nesting faces of the polytope such that all vertices are used once and only once. In that case, valid nesting configurations consist in specific permutations of the initial index sequence. As, for each cube in the tesseract, there are three possible pairs of square systems corresponding to parallel faces of the cube, there is a total of $4 * 3 * 3 = 36$ possible permutations such that each lower scale system contains only causal flows.

Among these 36 possibilities, 6 are dual solutions. For 6 others, which we call Primer Preserving Permutations (PPPs), the system formed by the primers of each lower-scale system is itself a face in the polytope. PPPs preserve the role of elements with index 2^p as being primers of one of the system in the configuration. Whereas the list of PPPs is easy to tabulate for a tesseract, a recursive algorithm can be used for larger values of n. Note that P_0 and P^* are both PPPs (see Fig. 5).

All configurations referred to in this section are made of four non-adjacent faces on the polytope, whose primers are related by a fifth upper-scale system. In the case of PPPs, the upper-scale system is itself a face in the polytope.

Dynamic Nesting Another way to define a nesting configuration is to construct it *on-the-fly*, by determining successively for each element placed in contrastive position, which of the possible implication systems it is more advantageous to relate it to. In this case, the cost function is used for each system hypothesis, to select the optimal one and disambiguate the antecedent function when several options are possible. Looking at Fig. 4, it appears that nodes $7, 11, 13, 14$ are contrastive in three different implication systems and 15 in 6 implication systems. Therefore, there exists $3^4 * 6 = 486$ distinct dynamic nesting configurations.

3 Optimization and Algorithmical Aspects

3.1 A Minimum Description Length Criterion

Given a sequence $X = x_0 \dots x_{l-1}$, the estimation process of the best PGLR, S^X, requires the definition of an optimality criterion embedding all the variables:

$$S^X = argmin_{P,G,R}\ \mathcal{F}(P, G, R|X)) \tag{2}$$

where P, G and R respectively denote the description of the polytope, the graph and the latent relations for sequence X.

Assuming that $\mathcal{F}$ is measuring the complexity of the sequence structure, S^X can be defined as the shortest description of the sequence. Therefore, searching for S^X can be seen as a Minimum Description Length (MDL) problem (20) and $\mathcal{F}$ can be understood as a function that evaluates the size of the "shortest" program needed to reconstruct the data (6). This is strongly related to the concept of Kolmogorov complexity, which has received increasing interest in the music computing community over the past years (11; 12; 13; 17).

The exact computation of S^X cannot be achieved and it is approximated in the following way:

– the description cost of P can be estimated as a function of the regularity of the polytope. In this work it is discarded because all polytopes are tesseracts.
– the description cost of G can be assumed to be constant for all configurations within a model class. It is related to the number of distinct possible graphs (DPG) in the PGLR.
– the cost ($\mathcal{F}_R$) of the relations associated with a given nested configuration:

$$R^X = argmin_R \; \{\mathcal{F}_R(R|G,X)\} \text{ with } \mathcal{F}_R(R|G,X) = \sum_{i=1}^{l-1} |r(\Phi_G(x_i), x_i)| \quad (3)$$

where Φ_G is the antecedent function associated to G and $|r(x,y)|$ is the cost of the relation between x and y.

3.2 Optimization Process

Given that the cost of P and G are assumed to be constant, the aim of the optimisation process is to estimate the set of latent relations.

In the case of TRs, the process is rather straightforward: a relation between two chords in a triad circle is unique.

Conversely, optimal transport provides multiple possibilities of connecting chords together. The exhaustive optimization over the whole sequence would require to consider all combinations of transports. However, to make the computation tractable, the process is divided in several simpler sub-problems as follows.

For the sequential model, the chord sequence is processed as groups of 4-chord progressions (fusing beforehand identical neighboring chords, for which the transport is trivially determined). Then the last chord of each group is related to the first chord of the next group by minimal transport.

For the static systemic models, each elementary problem corresponds to a square system to optimize. Upper-scale systems are optimized first and then each lower-scale system is estimated independently. This process is repeated for each possible configuration. Details can be found in (10), with two adjustments which do not significantly impact the performance but save a lot of computation load: (*i*) for square systems, the contrast relation is optimized aside from the other systemic relations, (*ii*) the set of static configurations is restricted to PPPs.

For the dynamic nested S&C model, each chord is considered successively in a chronological order. Those which are directly related to the primer (nodes 1, 2, 4 and 8) enable the estimation of the corresponding latent relation. Those who are in a single contrastive position (nodes $3, 5, 6, 9, 10, 12$) are used to complete the estimation of the corresponding systems. Some chords in contrastive position belong to several systems (nodes $7, 11, 13, 14$ to 3 systems and node 15 to 6 systems): in these cases, the system with minimal cost is chosen. The whole process therefore results in a graph which has been built dynamically by successive optimisations of square systems.

4 Experimental Validation

4.1 Methodology

Experimental Setups To assess the ability of the PGLR model to capture structural information in chord sequences, we have carried out a set of experiments on a corpus of 727×16 beat-synchronous chord sequences from the RWC POP dataset (5).

These experiments aim at evaluating the relevance of the PGLR model and at comparing different chord representations, types of models and optimization schemes.

The two types of chord representations presented in Sect. 2.1 (PCs and TRs) are considered in conjunction with optimal transport (for PCs and TRs) and triad circle relations (for TRs only). We compare the sequential bi-gram model (*Seq*) – a very common approach in MIR (15) – with different types of systemic models (*Sys* and *S&C*) as defined in terms of their antecedent functions in Table 1, as well as the dynamic approach (*Dyn*).

For the systemic models, three types of system optimization are considered:

- S_0 which corresponds to the static configuration P_0 (see Fig. 5, left);
- S^* which corresponds to the globally optimal PPP over the whole corpus which happens to be P^* (see Fig. 5, right);
- S^X: in this case, the optimal PPP is chosen a posteriori as the one that optimizes the description of X, which varies across all Xs.

Perplexity As there exists no ground truth as of the actual structure of a chord sequence, we compare the different models as regards their prediction ability. This is done by calculating for each model the *perplexity* (2), B, derived from the *negative log likelihood* (NLL), H. The aim is to measure how well an unseen sequence, $X = x_0 \dots x_{l-1}$, can be predicted by the model:

$$H(X) = -\frac{1}{l}\sum_{i=0}^{l-1} \log P(x_i|\Phi(x_i)) \tag{4}$$

with the convention $\phi(x_0) = x_0$ and $P(x_0|x_0) = P(x_0)$.

For the triad circle relations, $P(y|x)$ is estimated as the relative frequency of $r(x, y)$ (and $P(x_0)$ is set to $1/24$). Similarly, for a pitch class distance d, $P(y|x)$ is also estimated as the frequency of $d(x, y)$ (and here, $P(x_0) = 1/12$). The learning phase for r and d is done using a 2-fold cross-validation strategy: probabilities are estimated on one half of the corpus (even numbered songs) and used on the other half (odd numbered songs) to compute H and vice-versa.

For optimal transport, X is viewed as a set of simultaneous “voices”, X^k, and we compute H as the average voice NLL:

$$H(X) = \frac{1}{k}\sum_{k} H(X^k) \tag{5}$$

Table 2: Average perplexity obtained with 2-fold cross-validation for the different models on RWC POP. *DPG* stands for Distinct Possible Graphs.

	Triad Circle Rotation	Optimal Transport				DPG
		Chromatic		Harmonic		
	on TR	on PC	on TR	on PC	on TR	
Seq	8.00	3.32	3.58	4.11	4.50	1
Sys_0	8.88	3.43	3.68	4.32	4.72	1
Sys^*	7.62	3.12	3.11	3.86	4.23	1
Sys^X	5.78	2.66	2.73	3.18	3.41	6
$S\&C_0$	6.68	2.97	3.16	3.92	4.06	1
$S\&C^*$	5.35	2.60	2.71	3.39	3.56	1
$S\&C^X$	**4.63**	**2.39**	**2.48**	**2.99**	**3.12**	6
Dyn^X	4.82	2.55	**2.44**	4.29	4.32	486

where each term $H(X^k)$ can be computed horizontally, using Eq. 4.

Ultimately, the performance is reported in terms of *perplexity*, B, which can be understood as an estimation of the average branching factor in predicting the sequence knowing its PGLR structure:

$$B(X) = 2^{H(X)} \tag{6}$$

Note that, whereas PGLR is fundamentally optimized on the basis of a complexity criterion, its impact is evaluated in a probabilistic framework, so as to measure its capacity to compress the data information in a meaningful way.

4.2 Results

Table 2 summarizes the perplexity figures obtained for a variety of experimental setups, from which a number of observations can be made.

Benefit of Systemic Organizations Systemic models globally outperform the sequential one[1]: all perplexity values are lower, except for the basic Sys_0 configuration. In particular, the $S\&C^X$ model provides the most spectacular perplexity improvement for all types of chord representations and relations (at the expense of a very limited number of DPGs). Note that the P^* configuration provides a noticeable advantage over Seq and P_0 configurations. The last row of the table also shows that the dynamic nesting approach is an interesting alternative as it provides perplexity scores almost as favorable as $S\&C^X$.

Predictive Support of the Virtual Element The effectiveness of the virtual element in the S&C scheme is underlined by the systematic improvement

[1] This confirms preliminary results formerly obtained on a much smaller corpus of 45 chord sequences (10).

observed when shifting from Sys to $S\&C$ results. The virtual element, $\hat{x_3}$, in $S\&C$ appears globally as a better antecedent for x_3 than does the primer, x_0, in Sys. However, for about one third of test sequences Sys^X outperforms $S\&C^X$ (figure not reported in Table 2), in particular for *aaba* structures.

Triad Circle Relations vs Optimal Transport The performance of triadic circle relations (TCRs) is based on a global sequence entropy while the optimal transport (OT) approach is evaluated in terms of average "per voice" entropy. In particular, the maximal branching factor of TCRs is 24 instead of 12 for OT. Therefore, the two perplexity scores cannot be compared. However, both approaches show similar trends w.r.t. the relative model performance. This supports the hypothesis of a general benefit of the multiscale approach rather independently from the way the chord information and relations are encoded.

In Table 2, results are also provided for optimal transport on triadic reductions (TRs) treated as PC description. Here too, the relative performance levels across models show the same trends. Note that the perplexity on TRs is slightly higher because the average pitch class distance between triads tends to be larger than that between chords with 4 notes or more.

Harmonic vs Chromatic Transport In chromatic optimal transport, the distance is computed from the set of note displacements measured on a semitone scale. We also tested a harmonic distance by considering displacements on the circle of fifths. Results in Table 2 show that this globally degrades the performance. Conversely, there is no need to consider triad rotations on a chromatic circle, as this is formally equivalent to modeling systems on the circle of thirds.

5 Conclusions

Both from the conceptual and experimental viewpoints, the PGLR approach appears as an efficient way to model multiscale relations in music segments. It is expected to provide a useful framework for a number of tasks in automatic music processing, as well as offering an interesting tool for music analysis.

Given that its core principles are not specific to a particular type of musical information, the application of PGLR to other types of musical objects, such as melodic motives and rhythmic patterns is a rather natural extension, currently under investigation. Ongoing work also includes the extension of the PGLR model to a larger range of timescales (n-cubes) and to chord patterns of other lengths (using irregular polytopes, by truncating or duplicating vertices, edges or faces, as in (7)).

Bibliography

[1] Bimbot, F., Deruty, E., Sargent, G., Vincent, E.: System & Contrast : A Polymorphous Model of the Inner Organization of Structural Segments within Music Pieces. Music Perception 33, 631–661 (June 2016), former version published in 2012 as Research Report IRISA PI-1999, hal-01188244

[2] Brown, P.F., Pietra, V.J.D., Mercer, R.L., Pietra, S.A.D., Lai, J.C.: An estimate of an upper bound for the entropy of english. Computational Linguistics 18(1), 31–40 (1992)

[3] Cohn, R.: Audacious Euphony: Chromatic Harmony and the Triad's Second Nature. Oxford University Press (2011)

[4] Deruty, E., Bimbot, F., Van Wymeersch, B.: Methodological and musicological investigation of the System & Contrast model for musical form description. Research Report RR-8510, INRIA (2013), hal-00965914

[5] Goto, M., Hashiguchi, H., Nishimura, T., Oka, R.: RWC Music Database: Popular, Classical and Jazz Music Databases. In: ISMIR. vol. 2, pp. 287–288 (2002)

[6] Grunwald, P., Vitányi, P.: Shannon information and Kolmogorov complexity. arXiv preprint cs/0410002 (2004)

[7] Guichaoua, C.: Modèles de compression et critères de complexité pour la description et l'inférence de structure musicale. Ph.D. thesis, Rennes 1 (2017), directed by F. Bimbot, to appear

[8] Lerdahl, F., Jackendoff, R.: An overview of hierarchical structure in music. Music Perception: An Interdisciplinary Journal 1(2), 229–252 (1983)

[9] Lewin, D.: Some ideas about voice-leading between pcsets. Journal of Music Theory 42(1), 15–72 (1998)

[10] Louboutin, C., Bimbot, F.: Description of Chord Progressions by Minimal Transport Graphs Using the System & Contrast Model. In: ICMC. Utrecht, NL (2016)

[11] Louboutin, C., Meredith, D.: Using general-purpose compression algorithms for music analysis. Journal of New Music Research (2016)

[12] Mavromatis, P.: Minimum description length modelling of musical structure. Journal of Mathematics and Music 3(3), 117–136 (2009)

[13] Meredith, D.: Music analysis and Kolmogorov complexity. Proceedings of the 19th Colloquio d'Informatica Musicale (XIX CIM) (2012)

[14] Narmour, E.: The analysis and cognition of melodic complexity: The implication-realization model. University of Chicago Press (1992)

[15] Pearce, M.T.: The construction and evaluation of statistical models of melodic structure in music perception and composition. City University London (2005)

[16] Straus, J.N.: Uniformity, balance, and smoothness in atonal voice leading. Music Theory Spectrum 25(2), 305–352 (2003)

[17] Temperley, D.: Probabilistic models of melodic interval. Music Perception 32(1), 85–99 (2014)

[18] Tymoczko, D.: The geometry of musical chords. Science 313(5783), 72–74 (2006)

[19] Tymoczko, D.: Scale theory, serial theory and voice leading. Music Analysis 27(1), 1–49 (2008)

[20] Vitányi, P.M., Li, M.: Minimum description length induction, Bayesianism, and Kolmogorov complexity. IEEE Trans. Information Theory 46(2), 446–464 (2000)

[21] Weitzmann, C.F., Saslaw, J.K.: Two Monographs by Carl Friedrich Weitzmann: Part I:" The Augmented Triad"(1853). Theory and Practice 29, 133–228 (2004)